Base-10 (or, more generally, base-\(b\)), representation of integers

Almost all of school mathematics (including all the algorithms for arithmetic, etc.) relies on our base-10 system for representing numbers; understanding and being fluent in working with this system is a key part of school mathematics. (This system is also related in important ways to the ways we think about polynomials, as we’ll see later.) But how do you know that you can represent any integer this way? Prove the following proposition. (Hint: Use the division theorem and induction. You might also want to think about the fact that there is some \(\ell \in \mathbb{W}\) such that \(b^{\ell + 1} > m \geq b^\ell\).)

**Proposition.** If \(b \geq 2\) is an integer, then every positive integer \(m\) has an expression “in base \(b\)”: there are integers \(d_i\) with \(0 \leq d_i < b\) for all \(i\) and \(d_k \neq 0\) such that

\[
m = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_1 b + d_0.\]

Moreover, this is expression is unique.

The numbers \(d_k, d_{k-1}, \ldots, d_0\) are called the \(b\)-adic digits of \(m\).

Let \(m\) be a positive integer. Since \(b \geq 2\), there are powers of \(b\) larger than \(m\). We prove, by induction on \(k \geq 0\), that if \(b^k \leq m < b^{k+1}\), then \(m\) has an expression

\[
m = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_0.\]

The base case is \(k = 0\). If \(k = 0\), then \(1 = b^0 \leq m < b^1 = b\), and we can take \(d_0 = m\).

For the induction step, assume that the statement is true for all \(k < \ell\) and consider an \(m\) with \(b^k \leq m < b^{k+1}\). Applying the division algorithm, we see that there are integers \(q, r\) with \(0 \leq r < b^\ell\) such that \(m = q b^\ell + r\). Notice that \(q < b\), since otherwise we would have \(m \geq b^{\ell+1}\). Set \(d_\ell = q\). If \(r = 0\), we take \(d_{\ell-1}, \ldots, d_0 = 0\). Then \(m = d_\ell b^\ell + 0 b^{\ell-1} + \cdots + 0\) is an expression for \(m\) in base \(b\).

If \(r > 0\), we know \(r < b^\ell\) so the inductive hypothesis says that \(r\) has a n expression in base \(b\) in which the highest power of \(b\) with a nonzero coefficient is less than \(\ell\). Adding the term \(d_\ell b^\ell\) to this gives us an expression for \(m\) in base as required. This completes the induction proof of the existence of base \(b\) representations for all positive integers \(m\).

To prove the uniqueness, we also apply induction. First observe
that, since \(0 \leq d_i < b\) for all \(i\), then
\[
\sum_{i=0}^{k} d_i b^i \leq \sum_{i=0}^{k} (b - 1)b^i = \sum_{i=0}^{k} b^{i+1} - \sum_{i=0}^{k} b^i = b^{k+1} - 1 < b^{k+1}.
\]

Now we prove, by induction on \(k\) as before, that if \(b^k \leq m < b^{k+1}\), then the \(b\)-adic digits in the expression \(m = \sum_{i=0}^{k} d_i b^i\) are uniquely determined by \(m\).

The base case is \(k = 0\), so \(1 \leq m < b\). In that case, we must have \(m = d_0\) (and any higher coefficients being 0), so it’s clear that \(m\) uniquely determines the \(d_i\).

Now assume that we know the \(b\)-adic digits are uniquely determined if \(m < b^\ell\) and suppose we have two base \(b\) representations of a number \(m\) with \(b^\ell \leq m < b^{\ell+1}\), say \(m = \sum_{i=0}^{\ell-1} d_i b^i = \sum_{i=0}^{\ell-1} c_i b^i\), with \(0 \leq d_i, c_i < b\) for all \(i\).

If \(d_\ell \neq c_\ell\), we can assume that \(d_\ell > c_\ell\). Observe that \(\sum_{i=0}^{\ell-1} d_i b^i\) and \(\sum_{i=0}^{\ell-1} c_i b^i\) are both nonnegative numbers less than \(b^\ell\) by (\(*)\) above, so their difference must be strictly greater than \(-b^\ell\). So we have
\[
\sum_{i=0}^{\ell} d_i b^i - \sum_{i=0}^{\ell} c_i b^i = (d_\ell - c_\ell) b^\ell + \left(\sum_{i=0}^{\ell-1} d_i b^i - \sum_{i=0}^{\ell-1} c_i b^i\right) > -b^\ell.
\]

The left-hand side of this equation must be 0, since we are taking the difference between two base \(b\) representations of \(m\). But the right-hand side must be at least \(b^\ell - (b^\ell - 1) \geq 1\). So we must have \(d_\ell = c_\ell\).

Then consider the two representations of \(m - d_\ell b^\ell\) given by \(\sum_{i=0}^{\ell-1} d_i b^i\) and \(\sum_{i=0}^{\ell-1} c_i b^i\). The induction hypothesis says that \(d_i = c_i\) for \(i = 1, \ldots, \ell - 1\). So we have shown that \(d_i = c_i\) for all \(i = 1, \ldots, \ell\).

Do the following problems from the textbook

- Exercise 6.7: Now let us look for units in \(\mathbb{Z}[\sqrt{2}]\). Certainly, every nonzero number \(a + b\sqrt{2}\) has some real number as its multiplicative inverse, but this does not mean that \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\). All it tells us is that \(a + b\sqrt{2}\) is a unit in the ring of real numbers. The question is, does the real number that is the multiplicative inverse of \(a + b\sqrt{2}\) lie in \(\mathbb{Z}[\sqrt{2}]\)?

1. Suppose that \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\) and that its multiplicative inverse is \(c + d\sqrt{2}\). Show that \(ac + 2bd = 1\) and \(ad + bc = 0\).
We have \((a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}\).
If this is equal to 1, then we must have \(ac + 2bd = 1\) and \(ad + bc = 0\).

2. Using these equations, deduce that \(a - b\sqrt{2}\) is also a unit in \(\mathbb{Z}[\sqrt{2}]\), and that its inverse is \(c - d\sqrt{2}\).

Suppose \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\) and that its multiplicative inverse is \(c + d\sqrt{2}\). Then \((a - b\sqrt{2})(c - d\sqrt{2}) = (ac + 2(-b)(-d)) + (ad + bc)\sqrt{2}\). Since we know that \(ac + 2bd = 1\) and \(ad + bc = 0\), we can conclude that \((a - b\sqrt{2})(c - d\sqrt{2}) = 1 + 0\sqrt{2} = 1\), so \(a - b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\) and its inverse is \(c - d\sqrt{2}\).

3. Continuing with these numbers, show that the product
\[(a + b\sqrt{2})(a - b\sqrt{2})(c + d\sqrt{2})(c - d\sqrt{2})\]
equals 1 and deduce that \(a^2 - 2b^2\) must equal 1 or \(-1\).

Since this is a product of real numbers and we know that multiplication of real numbers is commutative, we can rearrange the order of factors to
\[
\left((a + b\sqrt{2})(c + d\sqrt{2})\right)\left((a - b\sqrt{2})(c - d\sqrt{2})\right) = 1 \cdot 1.
\]
Multiplying the first two factors and the last two factors, we get \(\left((a^2 - 2b^2) + 0\sqrt{2}\right)(c^2 - 2d^2) = 1\). But these two factors are integers, so either both are 1 or both are \(-1\). Thus, if \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\), we must have \(a^2 - 2b^2 = \pm 1\).

4. You have proved that if \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\), then \(a^2 - 2b^2 = \pm 1\).

As noted above.

5. Conversely, suppose that \(a\) and \(b\) are integers satisfying either \(a^2 - 2b^2 = 1\) or \(a^2 - 2b^2 = -1\). Prove that \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\). What is its multiplicative inverse?

Given \(a, b\) as above, consider \((a + b\sqrt{2})(a - b\sqrt{2}) = a^2 - 2b^2\). If this is 1, then the multiplicative inverse of \(a + b\sqrt{2}\) is clearly \(a - b\sqrt{2}\). Suppose instead that \(a^2 - 2b^2 = -1\). Then \((a + b\sqrt{2})(-a - b\sqrt{2}) = 1\) and we see the multiplicative inverse of \(a + b\sqrt{2}\) is \(-(a - b\sqrt{2}) = (-a) + b\sqrt{2}\).

6. Conclude that the units in \(\mathbb{Z}[\sqrt{2}]\) correspond to solutions to the Diophantine equations
\[x^2 - 2y^2 = 1; \quad x^2 - 2y^2 = -1.\]

We have shown in part 3 that if \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\), then \((a, b)\) is a solution to one of these Diophantine
equations. And in part 5, we showed that if \((a, b)\) is a solution to either of these Diophantine equations, then \(a + b\sqrt{2}\) is a unit in \(\mathbb{Z}[\sqrt{2}]\). So the units in \(\mathbb{Z}[\sqrt{2}]\) correspond exactly to the solutions to these Diophantine equations.

7. Observe that \((1, 1)\) is a solution to one of these equations, as is \((3, 2)\). Deduce that \(\sqrt{2} + 1\) is a unit, with inverse \(\sqrt{2} - 1\) and \(3 + 2\sqrt{2}\) is a unit with inverse \(3 - 2\sqrt{2}\).

This is just checking that \((1, 1)\) and \((3, 2)\) are solutions and then applying what we’ve proved in part 5.

8. Observe that \(3 + 2\sqrt{2}\) is just \((\sqrt{2} + 1)^2\) and the inverse of \(3 + 2\sqrt{2}\) is \((\sqrt{2} - 1)^2\). More generally, show that \((\sqrt{2} + 1)^n\) is a unit for every positive integer \(n\) by describing its inverse. Observe that we get in this way infinitely many units in \(\mathbb{Z}[\sqrt{2}]\).

Consider \((1 + \sqrt{2})n\). Using the associative and commutative properties of multiplication of real numbers, we can rewrite this as

\[
\left((1 + \sqrt{2})(-1 + \sqrt{2})\right)^n,
\]

which is clearly just \(1^n = 1\). So \((1 + \sqrt{2})\) is a unit, with inverse \((-1 + \sqrt{2})^n\). The elements \((1 + \sqrt{2})^{n_1}\) and \((1 + \sqrt{2})^{n_2}\) must be distinct if \(n_1 \neq n_2\); if not, let \(n_1 > n_2\) and we see that \((1 + \sqrt{2})^{n_1-n_2} = 1\). But that would make \((1 + \sqrt{2})^{n_1-n_2-1}\) the multiplicative inverse of \(1 + \sqrt{2}\). I’ll let you check that this is impossible. (Think about the binomial theorem, for example.) So all these powers of \(1 + \sqrt{2}\) are distinct units, and there are infinitely many units in \(\mathbb{Z}[\sqrt{2}]\).

9. Prove that if \(a\) is an even integer and \(b\) is an integer, the number \(a + b\sqrt{2}\) cannot be a unit in \(\mathbb{Z}[\sqrt{2}]\). (Is it possible for such a pair \((a, b)\) to satisfy the equation \(a^2 - 2b^2 = \pm 1\)?) Conclude that there are infinitely many numbers in \(\mathbb{Z}[\sqrt{2}]\) that are not units.

Suppose \(a = 2k\) is even. Then \(a^2 - 2b^2 = 4k^2 - 2b^2 = 2(2k^2 - b^2)\), so \(a^2 - 2b^2\) is even. But if \(a + b\sqrt{2}\) is a unit, we must have \(a - 2b^2 = \pm 1\), which is not even. Therefore, if \(a\) is even, \(a + b\sqrt{2}\) cannot be a unit in \(\mathbb{Z}[\sqrt{2}]\). Since there are infinitely many even integers, this shows there are infinitely many non-units in \(\mathbb{Z}[\sqrt{2}]\).
Exercise 6.8: In fact, we can show that in some sense the units are few and far between in the ring \( \mathbb{Z}[\sqrt{2}] \).

1. For a positive integer, how many numbers \( a + b\sqrt{2} \) are there in \( \mathbb{Z}[\sqrt{2}] \) with \(-N \leq a, b \leq N\)?

   There are \( 2N + 1 \) integer values of \( a \) with \(-N \leq a \leq N \), and (obviously) the same number of values of \( b \). Since \( a + b\sqrt{2} = c + d\sqrt{2} \) if and only if \( a = c \) and \( b = d \) (since \( \sqrt{2} \) is irrational, so no integer multiple of \( \sqrt{2} \) lies in \( \mathbb{Z} \)), there are \( (2N + 1)^2 \) elements in \( \mathbb{Z}[\sqrt{2}] \) with \(-N \leq a, b \leq N \).

2. Observing that for a given \( a \) there is at most one integer \( b \) that satisfies the Diophantine equation \( x^2 - 2y^2 = 1 \) and at most one integer \( b \) that satisfies the Diophantine equation \( s^2 - 2y^2 = -1 \), deduce an upper bound on the number of units \( a + b\sqrt{2} \) of \( \mathbb{Z}[\sqrt{2}] \) with \(-N \leq a, b \leq N \).

   The statement in the book is not correct. For instance, if we take \( a = 3 \), then both 2 and \(-2 \) give \( a^2 - 2b^2 = 1 \). (The statement is true if we insist that \( b \) be nonnegative, or if we change it to say “there are at most two integers \( b \)”, since \( a^2 - 2b^2 = 1 \) implies that \( b^2 = (a^2 - 1)/2 \) and, for fixed \( a \) this quadratic equation in \( b \) has at most two roots.) The analysis for \( a^2 - 2b^2 = -1 \) is similar.

   So for each \( a \in [-N, N] \), there are at most 4 values of \( b \) that satisfy \( a^2 - 2b^2 = \pm 1 \), and thus at most 4 possibilities for units in \( \mathbb{Z}[\sqrt{2}] \) of the form \( a + b\sqrt{2} \). Since there are \( 2N + 1 \) integer values of \( a \in [-N, N] \), this means that the number of units \( a + b\sqrt{2} \) of \( \mathbb{Z}[\sqrt{2}] \) with \(-N \leq a, b \leq N \) is at most \( 4(2N + 1) = 8N + 4 \).

3. Argue that the proportion of units among the numbers in \( \mathbb{Z}[\sqrt{2}] \) with \(-N \leq a, b \leq N \) is at most \( 1/N \) for large \( N \).

   We have shown that the proportion of units among the numbers in \( \mathbb{Z}[\sqrt{2}] \) with \(-N \leq a, b \leq N \) is at most

   \[
   \frac{4(2N + 1)}{(2N + 1)^2}.
   \]

   But this is equal to \( 4/2N + 1 < 2/N \).

   Note that, as \( N \) gets large, this goes to 0.

Exercise 6.10: Find the units in \( \mathbb{Z}[i] \).

1. Suppose that \( a + bi \) is a unit in \( \mathbb{Z}[i] \) and that its multiplicative inverse is \( c + di \). Show that \( ac - bd = 1 \) and \( ad + bc = 0 \). Using these equations, deduce that \( a - bi \) is also a unit in \( \mathbb{Z}[i] \) and that its inverse is \( c - di \).
We have
\[(a + bi)(c + di) = (ac - bd) + (ad + bc)i = 1 = 1 + 0i.\]

Since \(i\) is not a real number, no integer multiple of \(i\) can be a real number; this implies that we must have \(1 = ac - bd\) and \(0 = ad + bc\).

Supposing that \(a + bi\) is a unit with inverse \(c + di\), consider the product \((a - bi)(c - di) = (ac + bd) - (ad + bc)i\). We have shown that \(ac - bd = 1\) and \(ad + bc = 0\), so this product is equal to \(1 + 0i\). This shows that \(a - bi\) is a unit with inverse \(c - di\) as claimed.

2. Continuing with these numbers, show that the product
\[(a + bi)(a - bi)(c + di)(c - di)\]
equals \(1\) and deduce that \(a^2 + b^2 = 1\). Show that this means that the only possibilities for units in \(\mathbb{Z}[i]\) are 1, \(-1\), \(i\), and \(-i\), and then check that these four numbers are indeed units in \(\mathbb{Z}[i]\).

We know that \((a + bi)(c + di) = (a - bi)(c - di) = 1\). Using the commutativity of multiplication for complex numbers, this says that the desired product is 1. But observe that \((a + bi)(a - bi) = a^2 + b^2\) and \((c + di)(c - di) = c^2 + d^2\). So \((a^2 + b^2)(c^2 + d^2) = 1\). But this is a product of nonnegative integers (since \(a, b, c, d \in \mathbb{Z}\)), so if \(a + bi\) is a unit in \(\mathbb{Z}[i]\), we must have \(a^2 + b^2 = 1\). This means that only one of \(a, b\) can be nonzero and, whichever one is nonzero must be a square root of 1 in \(\mathbb{Z}\), hence equal to \(\pm 1\). This shows that the only possible units in \(\mathbb{Z}[i]\) are \(\pm 1, \pm i\).

Since 1 and \(-1\) are units in \(\mathbb{Z}\), they’re clearly units in \(\mathbb{Z}[i]\). And we have \(i(-i) = 1\), showing that \(i\) and \(-i\) are units as well.

- Exercise 7.3: Suppose \(u\) is a unit in a ring \(R\).

1. Recall that units can be canceled: If \(r\) and \(s\) are elements of \(R\) satisfying \(ur = us\), then \(r = s\). Observe that the logically equivalent contrapositive takes the following form: If \(r, s\) are elements of \(R\) satisfying \(r \neq s\), then \(ur \neq us\).

   There isn’t much to say here. The reason we can cancel \(u\) from \(ur = us\) is that we can multiply the equation \(ur = us\) by \(u^{-1}\), giving \((uu^{-1})r = (uu^{-1})s\). Since \(uu^{-1} = 1\), this says that \(r = s\).

2. Suppose that \(R\) has only finitely many elements and that \(r_1, \ldots, r_t\) is a complete list of them, with no repetitions. Deduce that \(ur_1, \ldots, ur_t\) is also a complete list of the elements of \(R\), with no repetitions. We might describe this result by saying that multiplication by \(u\) shuffles \(R\).
Since cancellation holds for units, \( ur_i = ur_j \) if and only if \( r_i = r_j \) if and only if \( i = j \). So \( ur_1, \ldots, ur_t \) is a list of \( t \) distinct elements of \( R \). But there are only \( t \) elements of \( R \) (and \( t \) is finite), so this means that \( ur_1, \ldots, ur_t \) is a complete list of the elements of \( R \) with no repetitions.

3. Give an example to show that in contrast, if \( a \) is an arbitrary nonzero element of \( R \), the set \( ar_1, \ldots, ar_t \) may not be a complete list of elements of \( R \). (Again note that \( a = 0 \) is a trivial example [of when \( ar_1, \ldots, ar_t \) is not a complete list]. You need only produce a single ring \( R \) and a single nonzero element \( a \) of \( R \) in order to demonstrate this. Thus, multiplication by an arbitrary element may not shuffle \( R \).

Let \( R = \mathbb{Z}_6 \) and \( a = [2] \). Then \([2][0] = [0], [2][1] = [2], [2][2] = [4], [2][3] = [0], [2][4] = [2], [2][5] = [4]\).

More abstractly, if \( a \) is not a unit, then there is no \( r_i \) such that \( ar_i = 1 \), so the multiplicative identity cannot appear in the list \( ar_1, \ldots, ar_t \).

• Exercise 7.9: Explain why Fermat’s theorem follows as a special case of Theorem 7.8. Then prove Theorem 7.8 following the outline below.

Fermat’s theorem says that, in \( \mathbb{Z}_p \) (\( p \) prime), every nonzero element \( u \) satisfies \( u^{p−1} = 1 \). (The “1” on the right side of the equals sign is really the congruence class \([1]\).) We know that the units in \( \mathbb{Z}_m \) are the classes \([a]\) with \((a, m) = 1\); in \( \mathbb{Z}_p \) these are all the nonzero classes. Since there are \( p \) elements in \( \mathbb{Z}_p \), there are \( p−1 \) nonzero classes. So Fermat’s theorem is a special case of Theorem 7.8.

1. Suppose that the ring \( R \) has precisely \( t \) units, \( u_1, \ldots, u_t \). Suppose also that \( u \) is a unit of \( R \). Thus, \( u \) is one of the \( u_i \), but we do not care which one. Show that \( uu_1, \ldots, uu_t \) is also a complete list of the \( t \) units of \( R \), possibly listed in a different order form the original list. (Thus, multiplication by \( u \) shuffles the set of units.

We know that \( uu_i = uu_j \) if and only if \( i = j \) (because units can be canceled, as above). So the list \( uu_1, \ldots, uu_t \) is a list of \( t \) distinct elements of \( R \). If we can show that all the \( uu_i \) are units, we’ll know that \( uu_1, \ldots, uu_t \) is a list of \( t \) distinct units in \( R \) and, since there are only \( t \) units in \( R \), it must be all the units in \( R \). But \( u \) and \( u_i \) are both units, with multiplicative inverses \( u^{-1} \) and \( u_i^{-1} \). So \( (uu_i)(u_i^{-1}u^{-1}) = u(u_iu_i^{-1})u^{-1} = u(1)u^{-1} = uu^{-1} = 1 \). So each \( uu_i \) is a unit.

2. Using the unit shuffle, prove that \( u_1u_2\ldots u_t = (uu_1)(uu_2)\ldots(uu_t) \).

The lists \( u_1, \ldots, u_t \) and \( uu_1, \ldots, uu_t \) contain exactly the same elements. Since multiplication in \( R \) is commutative, the products of the two lists will be the same.
3. Rearrange the factors in this equality and use cancellation to deduce that $u^t = 1$.

We have $(uu_1)(uu_2)\ldots(uu_t) = u^t(u_1\ldots u_t) = u_1\ldots u_t$.
Since the product of units is a unit, we can cancel the $u_1\ldots u_t$, and we get $u^t = 1$. 